

Arsenii Ashukha

[Home page](#) / [Google Scholar](#) / [GitHub](#)

e-mail: ars.ashuha@gmail.com tel: +79651782915

I'm a Research Scientist at Samsung AI Center Moscow. I (almost) received a PhD in Bayesian deep learning [@bayesgroup](#) so I can do math (mostly with Gaussian distributions) and make big overcomplicated DNNs work :)

My research interests are focused on crafting and training reliable and data-efficient machine learning models.

Prior to that, I was a part of Yandex.Research team (a Russian search giant), and did internships at Yandex, Rambler, and Center of Deep Learning and Bayesian Methods NRU HSE. I received a master's degree at Moscow Institute of Physics and Technology and Yandex School of Data Analysis (eq. to MSc programm run by Yandex). I did a bachelor's degree at Moscow State Technical University with a major in applied math and computer science.

EDUCATION

- 2017-2021 PhD in Computer Science, **Centre of Deep Learning National Research University HSE**. My PhD is focused on *applications and understanding of stochastic deep learning models*, Advisor: [Dmitry Vetrov](#).
- 2015-2017 MSc in Computer Science, **Moscow Institute of Physics and Technology** jointly with **Yandex School of Data Analysis**, where I worked with [Dmitry Vetrov](#) on learning sparse deep neural networks.
- 2011-2015 BSc in Computer Science **Bauman Moscow State Technical University** with a double major in applied math and cs. I worked on linguistically motivated topic models with [Natalia Loukachevitch](#).

PROFESSIONAL EXPERIENCE

- **Lead Engineer (eq. Research Scientist), Samsung AI Center** (2018 - Now):
The topics I worked on include: uncertainty estimation, generative models, learnable data augmentation, self-supervised learning, transfer learning for RL. Since 2021 I'm helping to manage the lab.
- **Research Scientist, Yandex Research** (2017 - 2018):
I worked on a group-level sparsification and uncertainty estimation. The method was applied to accelerate the feature extraction process for image retrieval.
- **Research Intern, Centre of Deep Learning and Bayesian Methods NRU HSE** (2016 - 2017):
I worked on sparsification of DNNs. This line of research led to Sparse Variational Dropout and started a new class of sparsification methods.
- **Machine Learning Engineer Intern, Yandex Music Deep Learning Group** (summer of 2016):
I worked on feature extraction techniques for music data with convolutional neural networks. I also developed an evaluation of learned representations. The representations were used in the content-based recommendation system for yandex music.
- **Machine Learning Engineer Intern, Rambler&Co** (May - Oct 2015):
Worked on demographic classification and recommendation systems. My responsibility included improving the quality and performance of classifiers, automatic feature extraction algorithms, and recommendation algorithms. Stack of technologies: Hadoop, Hive, Spark, XGboost, VW, gensim.

PUBLICATIONS

🍊 marks representative works

Google Scholar: scholar.google.com/citations?user=IU-kuP8AAAAJ

- Arsenii Ashukha, Alexander Lyzhov, Dmitry Molchanov, Dmitry Vetrov
🍊 *Pitfalls of In-Domain Uncertainty Estimation and Ensembling in Deep Learning*, ICLR 2020 [[arXiv](#)]
- Dmitry Molchanov*, Arsenii Ashukha, Dmitry Vetrov
🍊 *Variational Dropout Sparsifies Deep Neural Networks*, ICML 2017 [[arXiv](#)]
- Dmitry Molchanov, Alexander Lyzhov, Yuliya Molchanova, Arsenii Ashukha, Dmitry Vetrov
Greedy Policy Search: A Simple Baseline for Learnable Test-Time Augmentation, UAI 2020 [[arXiv](#)]
- Andrei Atanov, Alexandra Volokhova, Arsenii Ashukha, Ivan Sosnovik, Dmitry Vetrov
Semi-Conditional Normalizing Flows for Semi-Supervised Learning, INNF@ICML 2019 [[arXiv](#)]
- Evgenii Nikishin, Arsenii Ashukha, Dmitry Vetrov
Unsupervised Domain Adaptation with Shared Latent Dynamics for RL, BDL@NeurIPS 2019 [[arXiv](#)]
- Andrei Atanov, Arsenii Ashukha, Kirill Struminsky, Dmitry Vetrov, Max Welling
The Deep Weight Prior, ICLR 2019 [[arXiv](#)]
- Kirill Neklyudov, Dmitry Molchanov, Arsenii Ashukha, Dmitry Vetrov
Variance Networks: When Expectation Does Not Meet Your Expectations, ICLR 2019 [[arXiv](#)]
- Andrei Atanov, Arsenii Ashukha, Dmitry Molchanov, Kirill Neklyudov, Dmitry Vetrov
Uncertainty Estimation via Stochastic Batch Normalization, ICLR Workshop 2018 [[arXiv](#)]
- Kirill Neklyudov, Dmitry Molchanov, Arsenii Ashukha, Dmitry Vetrov
Structured Bayesian Pruning via Log-Normal Multiplicative Noise, NeurIPS 2017 [[arXiv](#)]

CODE

- Check out very short and simple and fun to make implementations of ML algorithms:
 - [Gradient Boosting](#)
 - [Real NVP](#)
 - [Quantile Regression DQN \(Distributional RL\)](#)
- Also, check out more solid implementations (at least they can do ImageNet):
 - [Multi-GPU SimCLRv1](#)
 - [Ensembles \(Deep ensembles, Snapshot ensembles, cSGLD, FGE, etc.\)](#)

REVIEWING

- Conferences:
 - International Conference on Machine Learning, ICML (2019, 2020 top-33% highest-scored reviewers)
 - Neural Information Processing Systems, NeurIPS 2019 (top-50% highest-scored reviewers)
 - International Conference on Learning Representations, ICLR (2020, 2021)
- Workshops:
 - ICML Workshop on Invertible Neural Networks (2019, invertibleworkshop.github.io)

TEACHING

- Supervisor of reading clubs on machine learning at HSE and Yandex school of data analysis (since 2017)
- Talks and practical sessions at **Deep | Bayes** Summer School on Bayesian Deep Learning (since 2017)
- Lecturer, Moscow Institute of Physics and Technology: I was a lecturer and manager of the deep learning brunch of a faculty-wide machine learning course ~60 students (ml-mipt.github.io). Also, I taught deep learning and practical sessions on cutting-edge ML algorithms on a facultative course "Data Mining in Action" ~ 200 students (<https://bit.ly/3eRLGYp>). The goal of this course is to make ML education available for everyone for free.

SUPERVISION:

- Alexander Lyzhov (moved to The Center on Long-Term Risk), Deep Neural Network Ensembles: Analysis and Approaches to Diversification (MSc, 2020)
- Andrei Atanov (PhD candidate at EPFL), Effective Learning of Deep Neural Networks Ensembles (BSc, 2018), Learning Deep Models with Small Data (MSc, 2020)
- Evgenii Nikishin (PhD candidate at Cornell), Stability Improvement and Knowledge Transfer in Deep Reinforcement Learning (MSc, 2019)

FRAMEWORKS & PROGRAMMING LANGUAGES:

- I'm fluent in Python and I use to code in C/C++, Go, language is not a problem after all.
- I'm also fluent with common python libs such as NumPy, Matplotlib, scikit-learn, etc.
- My primary deep learning framework at the moment is PyTorch. Prior to that, I had a decent experience with Theano+Lagange and TensorFlow.
- I have experience with MapReduce, Hadoop, Hive, and Spark.